1. Why inverse modeling and data assimilation in the Earth System sciences?

This question could be answered in two ways depending on whether one is starting form point of view of modeling or data. Both views are presented in the following.

1.1 Modeling in the Earth System Sciences

Physical theories in the Earth System sciences are designed to explain and possibly predict natural phenomena. The explanation by a theory is also a form of prediction as it states certain consequences for certain causes. Both, the explanation and prediction typically include quantitative representation of the natural system state. Quantitative assessment of the actual, true, state is fundamentally achievable only by measurements. The Geophysical theories are consequently designed to explain and predict the measurements.

The theories are most often expressed in a form mathematical relationships which define a model. The model in general represents governing physical laws and includes a set of quantities which entirely define the state by the model. The defining quantities are called control parameters. The control parameters are initial and boundary conditions, external forces and other physical quantities which define medium or environment for a process that is modeled.

The quantification of the system state by application of an assumed set of control parameters by the governing laws is called ***forward model*** of the system state or simulation of the measurements. Obviously, under conditions of well known governing laws and accurate quantification of the control parameters the forward model would produce accurate simulation of the measurements and would have ability to predict future states. It is common, however, that the governing physical laws are known but the control parameters values are not. This condition occurs in variety of models which are based on application of fundamental laws for macro scale phenomena such as conservation of energy and mass, propagation of energy through media and bulk energy and mass transformations. Examples of this type of model are found across the Geoscience disciplines such as in the Atmospheric Sciences, Oceanography, Biogeochemistry and Hydrology. What is typically not well known are initial and boundary conditions, some external forcing mechanisms and bulk properties of the medium in parameterized energy and mass transformations.

Because the model simulates the measurements, it is natural to ask whether there is a formal and objective way to use the measurements to infer the correct control parameter values for the model? This problem is called ***inverse problem*** or ***inverse modeling problem***. When there is a need to use the measurements to infer the control parameters frequently in a quasi continuous manner over time the inverse modeling problem is often referred to as ***data assimilation***. Thus, the first way of answering the question "Why inverse modeling and data assimilation in the Earth System sciences?" is

To objectively correct modeled state of the system or a component of it by using measurements, such that the model could be effectively used to analyze and predict the system.

The concept of inverse modeling and data assimilation for the purpose of improving model simulation and prediction has been used first in physical sciences in 17-th century in works by Euler, Lagrange and Laplace on calculating orbits of celestial bodies (Lewis et al, 2006). Gauss first formally described a method of data assimilation in his book on "Theoria Modus Corporum Coelestium" written in 1809. The data assimilation approach by Gauss is summarized in the following quote from the book:

" *If astronomical observations and other quantities on which the computation of orbits is based were absolutely correct, the elements aso, wheter deduced from three or four observations, would be strictly accurate, so for indeed as the motion is supposed to take place exactly according to the laws of Kepler, and, therefore, if other observations are were used, they might be confirmed but not corrected. But since all our observations and measurements are nothing more than approximations to the truth, the same must be true of all calculations resting on them, and the highest aim of all computations made concerning concrete phenomena must be approximate, as nearly as practicable, to the truth. But this can be accomplished in no other way than by suitable combination of more observations than the number absolutely required for the determination of the unknown quantities. This problem can only be properly undertaken when an approximate knowledge of the orbit has been already attained, which is afterwards to be corrected so as to satisfy all the observations in the most accurate manner possible."*

In chapter ?? we come back to the Gauss description of the data assimilation problem as it includes not only statement of the fundamental purpose of solving the problem but also key properties of the measurements and model. Since mid 19-th century the inverse modeling and data assimilation methodology continued to develop mostly in technical engineering applications where the primary problem was to either optimize a controlled system performance or to devise inference of signal over noise for measurements of dynamical systems (Jazwinski, 1970).

In the Geosciences the data assimilation for improving the model simulations was first explicitly used in 1980-es in Numerical Weather Prediction (NWP), where there is direct need to improve initial conditions to improve forecast skill (Daley, 1990; Kalnay, 2004). More recently, new research is being performed in which other than initial condition parameters are improved by the inverse modeling and data assimilation methods (Braswell et. al, 2005; Vukicevic et al., 2000).

1.2 Data in the Earth System sciences by data assimilation

First, the repeated data assimilation in the NWP naturally resulted in a time record of states which is used as data set for other than weather forecast. For example the weather analysis data are used in physical analysis of short term weather phenomena as well as in climate assessment and research on the climate processes (Trembert, ?; Reanalysis references). In late 1990-es the data assimilation emerged as necessary procedure in most other analysis of the Earth System states such as analysis of oceans, land surface, soil, atmospheric trace gasses and particulates, land hydrology and biogeochemistry. Every one of these analysis includes a model which control parameters are repeatedly corrected by the assimilation of measurements.

Examples of the Earth System data types containing temporally and spatially distributed physical quantities produced by some type of data assimilation are:

Ocean physical state data of

- velocity components
- pressure
- density
- temperature
- salinity

Ocean biological and chemical state data of

- concentration fields of nutrients
- plankton
- dissolved and particulate matter

Atmospheric physical state data of

- temperature
- pressure
- wind
- humidity
- cloud properties
- precipitation

Atmospheric chemical state and particulates data of

- concentration of trace gasses
- aerosols

Land surface state data of

- temperature
- moisture
- fluxes of gasses and energy
- optical and physical characteristics

Land biosphere data of

- land cover
- physiological characteristics of plants

Land hydrology data of

- hydraulic conductivity
- capillary pressure
- drainage coefficient in wetland areas
- leakage coefficient for river
- runoff

Soil state data of

- temperature
- moisture
- physical and chemical characteristics

These data are produced at operational environmental data centers and/or at sponsored national institutes and University research centers. In the USA the environmental data centers are sponsored or contained within all major national agencies: NOAA, NASA, DOE and DOD.

Thus, the second way of answering the question "Why inverse modeling and data assimilation in the Earth System sciences?" is

To produce quasi continuous fields of spatially and temporally distributed data of the Earth System, such that these data could be effectively used for the assessment of the system and for prediction.